

BIROn - Birkbeck Institutional Research Online

Atkins, Sarah and Roberts, C. (2016) Assessing institutional empathy in medical settings. *Journal of Applied Linguistics and Professional Practice* 13 (1), pp. 11-33. ISSN 2040–3658.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/26995/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>
contact lib-eprints@bbk.ac.uk.

or alternatively

Assessing institutional empathy in medical settings

Sarah Atkins and Celia Roberts

Abstract

The use of role-play in standardised medical assessments raises many problems, particularly around the measurement of interpersonal skills such as 'empathy'. Research on the use of simulations has tended to focus on quantified, psychometric assessments of their reliability and validity. However, communication, which often forms a central part of the assessment in medical simulations, is a difficult matter to address through this post hoc analytic method. A sociolinguistic approach to analysing real recordings of simulated medical assessments allows greater insight into communication and interpersonal skills. We use our research on medical licensing exams to illustrate some of the current questions about the examining of such skills through understanding the fine-grained detail of talk. Drawing on Goffman to critique simulated relationships and on Gumperz to analyse the role of differences in communicative style, and based on microanalysis of video-recorded role-players and candidates, we argue that the focus on interpersonal skills in standardised assessments amplifies the problem of using simulated empathy and requires additional interactional work. This focus on interpersonal skills in such assessments can lead to inequalities, since an unfair weight may be put upon candidates trained overseas. The paper concludes that the debate on 'language' in such exams and their standardisation in superdiverse societies needs to be reset.

KEYWORDS: ASSESSMENT; CLINICAL SKILLS; EMPATHY; PROSODY; SIMULATION;
STANDARDISATION; SUPERDIVERSITY

Contact author

Sarah Atkins, Centre for Research in Applied Linguistics, School of English, Trent Building, University of Nottingham, NG7 2RD, UK.
Email: Sarah.Atkins@nottingham.ac.uk

JALPP VOL 13.1-3 2016 @2018 11-33 <https://doi.org/10.1558/japl.31861>

©2018, EQUINOX PUBLISHING

 **equinox**
www.equinoxpub.com



This work is licensed under <https://creativecommons.org/licenses/by/4.0/>

1 Introduction

In a rather dark set of essays about empathy, the medical actor Leslie Jamison (2014) writes about how standardised, simulated consultations are used to assess medical students' and trainees' clinical skills. While these skills include such obvious categories as diagnostic eliciting and reasoning and clinical management, a significant component concerns what is variously called interpersonal effectiveness, compassion, affective behaviour and empathy. Jamison, as a standardised role-playing patient, is trained in the importance of checklist item 31: 'Voiced empathy for my situation/problem' where students 'have to say the right words to get credit for compassion' (Jamison 2014: 3). She writes:

I grow accustomed to comments that feel aggressive in their formulaic insistence: *that must be really hard* [to have a dying baby], *that must be really hard* [to be afraid that you'll have a seizure in the middle of the grocery store], *that must be really hard* [to carry in your womb the bacterial evidence of cheating on your husband].

(Jamison 2014: 4–5, italics and square brackets in original).

These glimpses into the experience of managing simulated empathy hold within them many conflicting positions: (1) the assumed value of standardised, simulated assessments; (2) the taken-for-granted importance of empathy, within the interpersonal skills area, both as a quality for all health professionals that *should* be assessed and one that *can* be assessed through simulated consultations; and (3) issues of fairness arising from these two. This article is structured around these three themes and refers to what is, at present, a small set of studies analysing simulated consultations from a linguistic and ethnographic perspective. We argue that the analysis of the fine-grained detail of the talk of standardised assessment and engaging with the larger concerns that shape the local interaction, such as professional discourses and institutional regimes, contribute to a growing debate about reliance on 'the unreal' (Greenhalgh 2014) in medical education and high-stakes assessments. In this paper, we draw on our research on the licensing exam of one professional body in the UK based on what are widely known as Objective Structural Clinical Examinations (OSCEs), and we suggest that the arguments we put forward are relevant for many such gatekeeping assessments.

2 Standardised, simulated consultations

Simulated consultations are widely used in Australia, Europe and North America, in undergraduate training and interim and final examinations (Bradley 2006; Cleland *et al.* 2009; Khan *et al.* 2013) as well as also for gaining membership of professional medical associations responsible for postgraduate curricula and licensing examinations (Swanson *et al.* 2013; First *et al.* 2013;

Khan *et al.* 2013). Their reputation in the medical world as the gold standard assessment is widely asserted (Boulet *et al.* 2009). In medical training, simulations provide opportunities for practice, reflection and focusing on certain skills, particularly technical skills (Korkiakangas *et al.* 2015), and so less importance is attached to authenticity. However, when used in exams, where they are known as standardised assessments or Objective Structured Clinical Examinations (OSCEs), they must be seen to be a fair and authentic reflection of real-life consulting skills. The widespread adoption of these standardised exams offers a replicable experience (Boulet *et al.* 2009; Cleland *et al.* 2009; Swanson and van der Vleuten 2013) readily able to be subjected to psychometric validation, thus aligning with the quantitative paradigms and models that are the hallmark of medicine.

The great bulk of research on simulated consultations has used psychometric tests to address reliability in terms of consistency and predictability (Bran-nick *et al.* 2011; Lievens and Sackett 2012) but fails to ask more fundamental questions about how like reality the interactions themselves are (Malhotra *et al.* 2009 being an exception), how power relations in the exam setting play out in the interaction (Atkins, forthcoming) and to what extent such simulations may disadvantage certain groups. Sociolinguistics and linguistic ethnography, in looking at the fine-grain detail of interactions and situating these within professional and institutional practices, can begin to answer these questions (Roberts *et al.* 2003; Roberts *et al.* 2014; Seale *et al.* 2007; De la Croix and Skelton 2009, 2013; O’Grady and Candlin 2013; Niements 2013).

In medical education, Niements (2013) looks at role-plays of interpreter-mediated consultations, describing how they ‘cannot reproduce the orientations of real interactions. [...] [W]hat is authentic to those users when they “live” a specific situation cannot be authentic to trainers/trainees when they play it’ (Niements 2013: 317). De la Croix and Skelton (2009) similarly find that the reality of the medical encounter is difficult to ‘reproduce’ in a simulation. They make a detailed sociolinguistic study on interactional power in medical simulations in an evidence-based, corpus linguistic analysis of undergraduate OSCEs. They find role-players talk and interrupt more than candidates, citing this as evidence of the conversational dominance role-players exert in a way that a patient in a real encounter tends not to do. However, there is still considerable scope for better understanding the asymmetric participation structures that can be established in simulated encounters and how these are instantiated in the moment-by-moment interaction.

Within medical research, any doubts about OSCE-type exams are often trumped by the psychometric findings on reliability and most types of validity. How simulated consultations are experienced differently from real ones and how this, in turn, may have a differential impact on certain groups of

candidates is masked by the methodologies of standardisation and statistical analysis and their disengagement from the local and the contextual. In addition to the psychometric evidence, these exams are defended on two somewhat contradictory grounds: that the exams are a proxy for the real – a good enough mimic of reality (so it does not matter that they are a set performance) – and that the simulated behaviour they require is appropriate since consultations are a performance anyway (so in some sense real). Goffman's extensive discussions of how we experience face-to-face interaction disturb both of these assumptions.

Goffman (2005) argues that the sense of feeling an activity is real depends upon our sense of self as we relate to others. Each interaction creates and reinforces a shared reality to keep the relationship going. It is through interactional frames, Goffman argues, that interactants feel real to each other and evaluate each other as they display their understanding of situated intent. But the simulated consultation is a complex encounter where multiple frames are at play (Goffman 1974: 156–200). For example, in an OSCE-style exam, the frame of showing empathy to a role-playing patient is nested in a frame of displaying competence to an examiner, which in turn is nested in the institutional frame of the overall assessment process. So what matters is not how emotionally and sincerely connected the candidate feels to the role-player but how far they are seen as 'empathic' by the examiner. The real quality of empathy is assessed through the unreal – the display of empathy for institutional assessment (Seale *et al.* 2007; De la Croix and Skelton 2013).

While one defence of exams using standardised patients is that they can be real enough – their simulated character is not attended to – they are also defended on the grounds that they are a performance and that all consultations are, in effect, 'performed'. This is a common response to those candidates who voice concern that OSCE examinations test acting skills as much as clinical. Again, Goffman helps with unpacking the subtleties of what is meant by performance and acting. The 'frontstage' performance (Goffman 1959) required, for example, of a doctor, teacher or waiter involves constraints on behaviour. And this behaviour requires a heightened performance when monitored and assessed for institutional purposes. But these frontstage and institutional performances are different from a simulated performance, where the event is acted. The actor has to be convincing within the terms of the play (in this case the simulated consultation), while monitoring their own behaviour and so sustaining a 'dual consciousness' (Konijin 2000). The trainee/candidate therefore has to work hard to create a synthetic reality – one that convinces the audience/observer – but not one that is real to candidates in terms of consequences for patients. Whereas the role-player's task is to be

authentic within the terms of the simulation/drama, the candidate's task is to display themselves as they would in a real consultation while also monitoring their conduct *vis-à-vis* the examiner and maintaining the illusion. This is hard communicative work (Thomassen 2009).

An additional complexity in the drama of the exam is some shift in power from the doctor to the role-player (De la Croix and Skelton 2009). Textbooks and training on communication and interpersonal skills assume an asymmetry, as Jamison's '*that must be really hard*' illustrates, with the trainee/candidate doctor gracing the 'patient' with their understanding. Being 'empathic' to someone who is in a more powerful interactional position than oneself is harder than it is to someone relatively less powerful, and puts additional demands on acting skills; and role-players' power is made even more difficult to manage, since while their behaviour is ostensibly standardised, precisely how they react and interact depends upon their instinctive resources, meaning that it is near impossible to standardise interaction completely. Role-players are instructed to react in specified ways to the candidate's behaviour but exactly how they do this is up to them (Atkins *et al.* 2016). It is important to stress that there is no simple inversion from powerful doctor (candidate) to powerful patient (role-player). Rather, it is that the role-player has much more freedom in how the interaction unfolds and the *potential* to claim power at that micro-level. Sometimes they are actually very compliant (even overly so), but (very often with candidates who are nervous or make a slight mistake) the role player has much more freedom to claim some power – withhold responses, highlight mistakes and so on.

A close look at the simulated consultation reveals a highly complex and hybrid activity. As well as multiple roles and identities at play, considerable interactional work has to be carried out to sustain the illusion of reality and to avoid or cover up any overt frame-breaking which would undermine this illusion (Seale *et al.* 2007: 181). The fact that not just the role-players, but also the candidate doctors themselves, are required to act is avoided in the discourse of the exams. This is, though, voiced backstage by some examiners, as in this feedback session looking at exam video clips:

I wonder whether he was having trouble acting this particular consultation. He just seemed a little bit kind of remote from him. I just wondered if he was having an acting problem this candidate

(examiner feedback from Roberts *et al.* 2014: Appendix D 47–54).

The experience of being in a role-play is fundamentally different from a real-life consultation (Niements 2013), and it is in the very elastic notion of 'interpersonal' that candidates' capacity to sustain the illusion of reality is most crucial and most vulnerable to negative criticism.

3 Patient-centredness and interpersonal skills

The major discursive shift from doctor-centred to patient-centred consulting over the last 50 years (Balint 1970; Stewart 2001) has brought with it changes in how to relate and communicate with patients. While the asymmetric relationship between doctor and patient has remained largely intact in contemporary healthcare interactions, the discourses of patient-centredness, and so the importance of interpersonal skills, have become thoroughly embedded in training and assessment. It is argued that interpersonal skills (IPS) should be rated alongside and as an equal partner with the bio-psychosocial approach (Howie *et al.* 2004).

This discursive shift is part of the more seismic movement in how individuals and relationships are conceptualised. This is the shift to the neo-liberal focus on the self and soft skills in which the social and cultural self becomes the centre of attention in assessment, as much as or even rather than acquired technical expertise (Grugulis and Vincent 2009). So, it is no surprise that the soft skills of interpersonal effectiveness and communication generally (and the distinction between them is not always clear) and ‘clinical empathy’ (Halpern 2003) in particular, appear in virtually all models of consultation and are taught as a matter of course (Neighbour 1987; Kurtz *et al.* 1998; Hojat *et al.* 2004; Bonvicini *et al.* 2009).

Most research and assessment of interpersonal and communication skills has been modelled on medical sciences. Such modelling reduces the complexity of these phenomena to patterns and procedures that can be statistically analysed within the same psychometric paradigm mentioned above (Skelton 2005; Iedema *et al.* 2006). This has led, despite some concerns (Campion *et al.* 2002; Howie *et al.* 2004), to a relatively unproblematic stance on the assessment of interpersonal effectiveness. However, sociolinguistic analysis has identified several overarching problems which suggest that the interpersonal is not easy to pin down. Firstly, IPS cannot be readily categorised as a separate domain, since the *manner* in which all aspects of the simulated consultation are carried out contributes to interpersonal evaluation (Roberts *et al.* 2014). The interpersonal leaks into everything. Secondly, the assessment tools are conceived as universal, and not relative to particular social groups: ‘[M]uch work has still to be done to make performance measurement into a culture-sensitive and equitable science’ (Howie *et al.* 2004: 464). Thirdly, despite their ‘objective’ naming, the interpersonal in such exams is, necessarily, subjectively assessed and, in settings of intense social evaluation, small differences and difficulties are amplified (Gumperz 1982).

Arguably, just as the interpersonal is the most problematic component to assess in standardised, simulated consultations, so the concept of empathy is the most untameable element within the interpersonal domain, intensifying the

problems outlined above in assessing IPS. The notion of empathy is widely taken for granted as part of good doctoring in mainstream medicine. In textbooks and training on consultation and communication skills, trainees are encouraged to voice empathy just as role-players and examiners are expected to look out for and acknowledge it, as widely used empathy models attest (Suchman *et al.* 1997). A typical example of modelling empathy is found in the Cambridge-Calgary guidelines on good clinical communication: ‘Work out exact phrases which demonstrate empathy in specific situations’ (Kurtz *et al.* 1998: 134). It is hardly surprising, therefore, that candidates voice standard empathy utterances, given the direction to perform these ‘phrases’ from the medical communication literature. However, from within sociology and ethics the notion of empathy and its ‘emotional labour’ (Larson and Yao 2005) is contested from three different perspectives:

- is it a quality that is essential for good doctoring?
- if it is, how can it be best understood and assessed?
- if it is to be assessed in OSCE-type exams, can simulated empathy be assessed as real empathy?

The assumption that empathy is a moral requirement for the practice of good medicine has begun to be critiqued. Bouma (2008) argues that there are important ethical values which trump empathy, such as patient autonomy, and that doctors can be caring, ethical professionals without showing ‘true’ empathy, defined as sharing patients’ emotions and being moved by their suffering (Halpern 2003). Similarly, it is suggested that doctors cannot be the superior moral beings that ‘true’ empathy requires, and recommend ‘*étiquette*’ – that is, being clear and courteous – as an acceptable goal: ‘[W]e need to make far more moderate claims about what is being taught, and how and why’ (Smajdor *et al.* 2011: 383). But despite such criticisms of the place of empathy in doctoring, it remains a dominant theme in the discourses and practices of training and assessment.

While there is some acknowledgement in the medical education literature that empathy is hard to define and measure (Hojat *et al.* 2004; Stepien and Baernstein 2006; Marshall and Bleakley 2009; Pedersen 2009), the fundamental problem remains the difficulty in objectively judging aspects of IPS such as empathy, rapport or sincerity from the outside. Empathy is an inner state, experienced (or not) only by someone to whom it is directed (Seale *et al.* 2007). Only those who are on the receiving end of intended empathy or its lack can say how they felt about that moment of the interaction.

In OSCE-type exams, this inner state cannot be judged, since the role-player is simulating their pain and their concerns. All that can be judged is the candidates’ display of empathy and, as Jamison points out, empathy and

compassion must be ‘voiced’ (Jamison 2014: 3–4) so that it is explicitly markable by the examiner. The display of empathy phrases also raises questions about whether soft skills such as empathy and rapport are, in real life, always verbalised explicitly or whether such skills are often conveyed much more indirectly and subliminally. A corpus linguistic study of how health professionals show compassion (Brown *et al.* 2006) shows that explicit compassion phrases are not frequently used.

In our study of an OSCE-type exam for licensing health professionals, successful candidates manage this displayed empathy well and indeed use empathy phrases more frequently than unsuccessful candidates (Roberts *et al.* 2014). And yet in examiner feedback sessions, where examiners were asked to comment on clips from video-recorded candidates, there were frequent criticisms that some candidates sounded formulaic and insincere. For example: ‘It seems just very formulaic and a lot of it seems learned: “I understand why you would be worried”’ (examiner feedback from Roberts *et al.* 2014: Appendix D, 41). Clearly the frequency of these phrases was not the issue, as high-performing candidates used them more often and they index the patient-centred model that frames such exams. What is perceived as ‘trained empathy’ (Roberts and Sarangi 2003) stems from both the design of the empathy phrases and how they are delivered. Successful candidates customised their voiced empathy, redesigning these phrases by adding small discourse markers and conversational phatics to relax these standardised phrases, as in the example below (here with punctuation added for clarity): ‘Do you know what – I really do understand that.’ So, in order to meet the requirements of the exam and yet not sound formulaic, candidates have to do additional communicative work. As well as the ‘double consciousness’ (Konijin 2000) of both acting authentically within the terms of the drama (assessment within a patient-centred model) and monitoring one’s performance, there is a third element – voicing ‘empathy’ laminated with ‘sincerity’ linguistic tokens; what we might call a ‘triple consciousness’ is at work.

Candidates also have to deal with *how* they sound in context. We draw heavily on Gumperz’s early work (Gumperz 1982: 100–129) on prosody and how differences can lead to misevaluation, misunderstanding and gaps between speaker intention and pragmatic assessment. And it is now possible to benefit from software that allows us to analyse the waveforms and pitch of each line of a speaker’s talk (MacWhinney 2000; ELAN 2017). So we can examine single utterances to see how different aspects of prosody (pitch, volume, rhythm and intonation contours) work together to give off contextual cues which, in turn, may form a particular impression. While the negative effects of different accents (including aspects of prosody) in assessing overseas doctors have been generally discussed (Hoekje 2011), the microanalysis of prosodic features indicates how complex and nuanced these different features

can be and how difficult to notice and understand unless recorded and subject to sociolinguistic analysis. In addition, the power relations in simulated consultations are likely to amplify negative attitudes. Although in real medical encounters such attitudes might be mitigated because of the relative high status of doctors, as Rubin *et al.* (1997) suggest, in the simulation, where the power dynamic is shifted, the same allowances for differences in accent and prosody may well not occur. The following examples look at prosody for two candidates and the assessments that accompanied their overall performance.

The first example, in Extract 1, is a case in which the role-player patient (RPL) has requested a medication not ordinarily provided by the UK National Health Service. The candidate (CAN) achieves full marks across all domains. This extract begins six minutes into the consultation, when the patient returns to the topic of his requested prescription (see Appendix for transcription conventions).

Extract 1

- 481 RPL: it does seem very unreasonable an- ≈
 482 CAN: +≈ yeah
 483 RPL: whereas if you
 484 almost if you cross the road they say oh that's fine
 485 CAN: yeah
 → 486 no i do do you know what i really do understand that
 487 i think particularly if you've got friends who
 489 have been um (.)
 490 [been allowed to have it it's]
 491 RPL: [w- well yes i was er yes i]
 492 [exactly]
 493 CAN: [it's] frustrating

The role-player simulates that he is disgruntled and in order to display 'empathy' with him and his frustrations the candidate uses a typical phrase at line 486. Here we can analyse the 'prosodic contour' with which she delivers the utterance. The rise and fall in volume and pitch are shown in Figure 1.

In local British English, information units are produced and processed in smooth prosodic 'envelope contours', with the volume and the pitch register generally following one another: a chunk of information is given within one single prosodic contour, the pitch typically going down at the end, as this candidate's smoothly does. The emphasis at the opening of the utterance, which is a little higher and louder than the rest, triggers a marked affect. After the false start at line 486 'No I do' the candidate says 'do you know what – I really do understand that' with a high pitch and raised volume at the onset of the main utterance 'do'. This whole line is enclosed in this one smooth enveloping affective contour, without a pause, and is rounded off at the end with a drop in

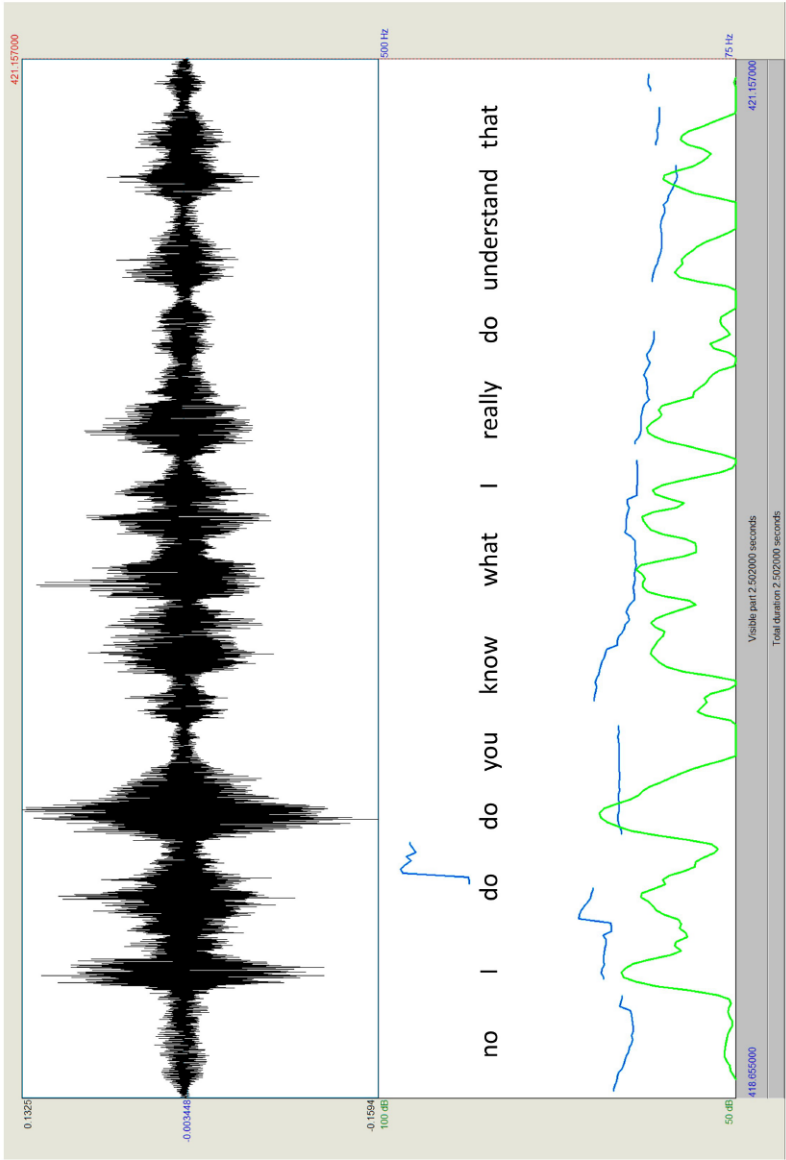


Figure 1: Prosodic and volume contour of an utterance by a successful candidate (lower line shows the rise and fall in volume; upper line shows the rise and fall in pitch)

pitch. She conveys expressiveness with the higher tone and makes the whole package sound smoothly and sincerely delivered.

She also customises her ‘understanding’ phrase, to make it sound less formulaic and more sincere. Firstly, she is able to adapt the wording a little, so that it sounds less like one of the rote-learned phrases that were commented on by examiners: ‘No I do – do you know what – I really do understand that’. Particularly the small emphasis marker ‘really’ and the more conversational focusing and framing prefix – ‘do you know what’ – would seem to stress her sincerity. She also follows up her expression with a longer account of why she can understand his frustrations (lines 487–490), at a point in the consultation when she can take up a little more time. In combination, these multi-channelled features of communication produce a highly rated performance. And although such small moments do not a whole consultation make, cumulatively the voiced empathy and the work done to mitigate any formulaic dangers – much of it below the level of consciousness as the examiner listens in – is a masterful exemplum of the ‘language game’ which constitutes standardised examinations.

4 Issues of fairness

The degree to which simulated consultations, and specifically the simulation of interpersonal skills, are valid means of assessment is of critical concern to any institution using them. An equally pressing concern is one of fairness in contexts of superdiversity (Vertovec 2007), particularly when there is evidence that ethnic minority examination candidates in UK, many of whom are international medical graduates (IMGs), are faring less well than UK-trained colleagues (McManus *et al.* 2013: 8).

In particular, our research on an OSCE-style licensing exam in the UK suggests that the assessment of simulated empathy may be a contributing factor in the low success rates of IMGs. In other words, the design of such exams can put a burden on some candidates that others do not feel. In Bourdieu’s (1977) terms, they are like fish who feel the constant weight of the water they swim in. Many features of the ‘hidden curriculum’ in postgraduate professional assessment can produce this weight. They cannot be fully discussed in this article (see Roberts *et al.* 2014), but they include the talk-heavy character of the exam (candidates on average talked for a greater amount of the floor time, approximately 68% in simulated consultation, in contrast to about 60% in real consultations) and its relatively decontextualised environment, with, for example, no shaping role of the computer in the surgery (Swingelhurst *et al.* 2012). In the specific context of empathy assessment, it is the difference between how UK medical graduates and IMGs use and deliver empathy phrases. As mentioned above, UK graduates can customise and conversationalise these phrases to

inoculate them against criticisms of formulaic consulting. However, IMGs, whose consulting practice overseas was unlikely to focus on such patient-centred manoeuvres, not only had to learn appropriate empathy utterances but also then learn that such phrases had to be redesigned.

More profoundly, the prosodic systems used in the varieties of English spoken by most IMGs differ from those of UK graduates (although of course many of these candidates may be increasingly influenced by the local English they hear around them). Again, this analysis draws on Gumperz's work on prosody and contextualisation cues (Gumperz 1982: 173–186) and how differences can lead to misevaluation and misunderstanding and gaps between speaker intention and pragmatic assessment. In contrast to the prosodic analysis of the successful UK graduate discussed above, the following case is of a more borderline candidate, who still passes the assessment but with a much lower score, achieving his lowest mark in the IPS domain.

The extract below comes from a complex case concerning a child protection issue, in which the role-playing mother (RPL) is concerned that her partner's brother is 'cuddling' their son. The candidate (CAN), of South Asian origin, does badly in IPS, achieving only 1 out of 3 and the examiners tick the following feedback comments: 'Does not appear to develop rapport or show sensitivity for the patient's feelings'; 'Does not make adequate use of verbal and non-verbal cues, poor active listening skills'.

The particular expression of understanding analysed in Extract 2 comes at lines 100–102, just after the mother has described her concerns.

Extract 2

- 93 RPL: and he said that um (1.4) dave who's the tom's brother cuddling him
- 94 CAN: right
- 95 RPL: um and had
- 96 said that it was their little secret
- 97 and not to tell anybody
- 98 CAN: mmm
- 99 PPP: (1.4)
- 100 CAN: right i can see why you are concerned
- 101 and er it's not something to
- 102 to be taken lightly and i can see where you're coming from
- 103 RPL: mmm
- 104 CAN: ok
- 105 and er
- 106 what happened after that did you
- 107 have a chance to speak to er

We can see lines 100–102 perform the similar interpersonal function of voicing understanding for the patient's situation as in Extract 1, albeit in a very different exam case. The three-part turn, reformulating the original phrase twice, suggests very strongly that this candidate both knows about displaying explicit empathy and intends to show that he is 'empathic' within the terms of the exam. The words he uses are very similar to the stock phrases all candidates use. If we look at the prosodic patterning, though, it is different from the candidate in Extract 1 and potentially causes some difficulties in how the expressions are received. His pitch is much lower, but it also varies less, not rising and falling as the volume of his talk does (Figure 2).

This candidate draws on aspects of North Indian languages such as Hindi in his use of volume, pitch and intonation. Word order is more flexible in languages like Hindi and a lot of the work that is done in English with stress and intonation is done by juxtaposing words and phrases (Gumperz 1982: 119–129). These prosodic patterns are quite common in this system of English influenced by North Indian languages where small units of talk are juxtaposed: 'right / I can see why you are concerned / and er / it's not something to to / be taken lightly / and I can see where you're coming from.' Each unit uses the same melody, so it sounds like a list and potentially rather formulaic to some ears (Gumperz 1982: 149). And while the volume is raised at the beginning of each information package, the pitch remains quite low and flat. In standard or local British English this could sound uncaring but in Hindi and in English influenced by North Indian languages low pitch is conventionally a marker of respect or conveying bad news (Gumperz 1982: 184).

Findings around the intonation of these token expressions are complex to discuss in the bigger picture of clinical skills assessment. Interpretation of meaning and attitude through prosody is largely unconscious and automatic, a result of learning to interact with others who use features 'like us'. This is clearly not something that is explicitly assessed, but nevertheless may impact on how utterances are understood and judged. In an exam where conventional interpersonal expressions pervade much of the talk, candidates' prosodic delivery is an important part of whether they stand out as 'formulaic', despite only marginal differences in the words used. It would seem from the grade and comments given to this candidate that such small differences are amplified under the intense gaze of the examiner and that this candidate is sanctioned for a lack of 'rapport'.

In feedback sessions, where examiners were asked to comment on segments of video-recorded candidate performances, many of the IMGs were perceived as formulaic even though the empathy phrases were similar to those of UK graduates. Formulaic-sounding phrases also led to wider judgements

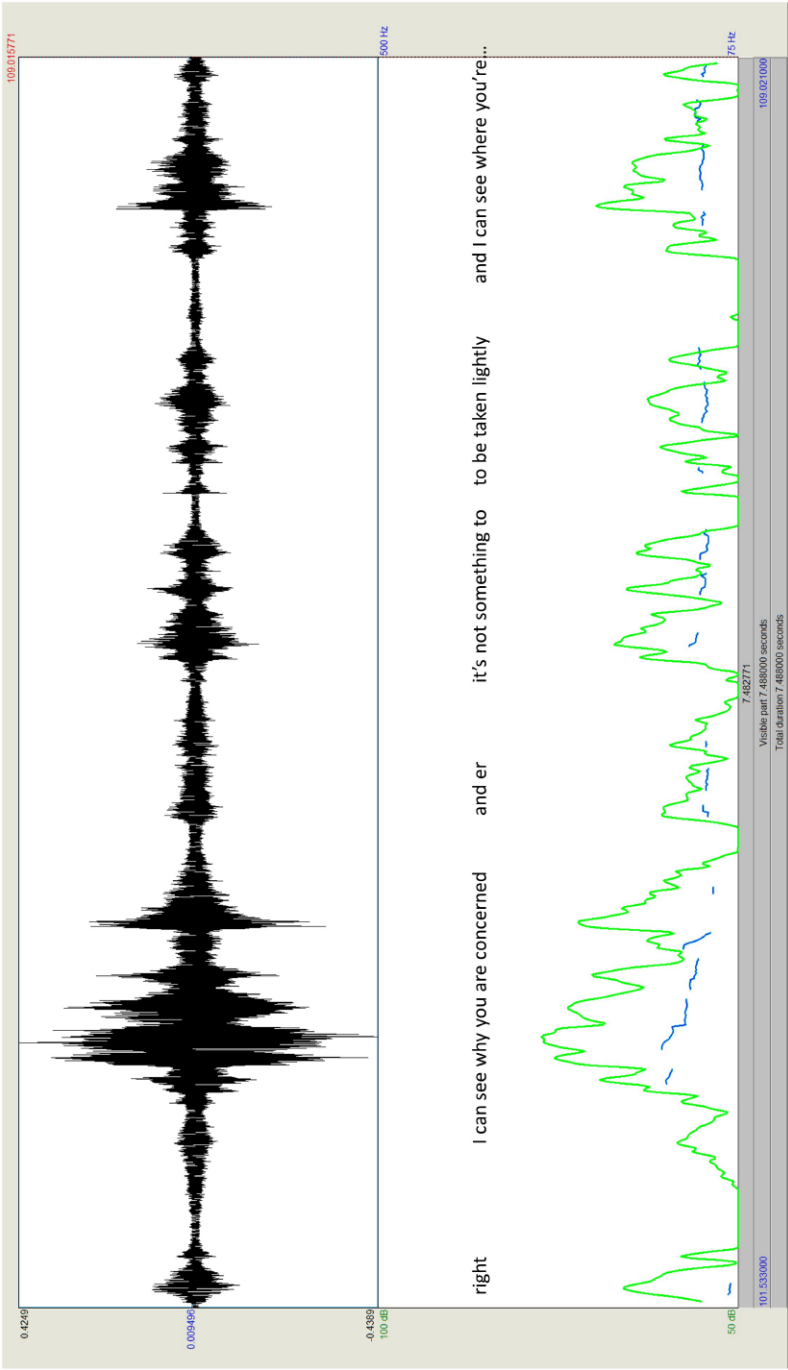


Figure 2: Prosodic and volume contour of an utterance by an unsuccessful candidate (upper line shows the rise and fall in volume; lower line shows the rise and fall in pitch)

of manner and attitude such as ‘not engaging’, ‘not interested enough’, ‘a bit Olympian’. So perceived differences can rapidly lead to ‘interpretive overdrive’ as Blommaert and Rampton discuss:

For much of the time, most of the resources materialised in any communicative action are unnoticed and taken for granted, but it only takes a slight deviation from habitual and expected practice to send recipients into interpretive overdrive, wondering what’s going on when a sound, a word, a grammatical pattern, a discourse move, or bodily movement doesn’t quite fit. There is considerable scope for variation in the norms that individuals orient to, which affects the kinds of things they notice as discrepant, and there can also be huge variety in the situated indexical interpretations that they bring to bear (‘good’ or ‘bad’, ‘right’ or ‘wrong’, ‘art’ or ‘error’, ‘call it out’ or ‘let it pass’, ‘indicative or typical of this or that’).

(Blommaert and Rampton 2016: 37)

Such perceived differences can affect both overall emotional tone – whether the candidate sounds warm, involved, responsive and so on – and overall behavioural smoothness, i.e. whether the interaction progresses without jarring or uncomfortable moments or not (Erickson and Shultz 1982: 169–173). If both role-player and examiner share normative expectations about communicative resources but these are not shared with candidates – even at the most micro-prosodic level – the whole encounter may seem to have gone awry. This may lead to candidates being assessed as lacking compassion or disengaged, and doubts about how they can be trusted to be a caring professional are raised (see also O’Grady and Candlin 2013).

In a superdiverse context, where we are increasingly exposed to one another’s different communicative styles – and no more so than in the health service – the ability to produce these phrases in a manner which sounds sincere to an over-hearing examiner, in what is a simulated context, may not be an especial priority as a doctoring skill. Indeed, in superdiverse contexts, there is a strong argument for rethinking what the interactional environment of consultations looks like and whether local British normative expectations of expressing compassion or empathy can remain the dominant style of assessment. Deciding whether talk is empathic or not depends upon whether the observing examiner considers that the candidate’s behaviour accords with what they consider empathy to be. The objectivity of standardised assessments is difficult to sustain when judgements such as ‘empathic’ depend upon an examiner’s unavoidably subjective assessment of candidates’ simulated behaviour towards simulated patients.

If patients also shared the same communicative expectations, there would be some case for arguing that all candidates should be able to draw on the same set of resources. However, the superdiverse patient populations which are characteristic of so many urban areas require a level of communicative flexibility that is not designed into standardised exams – indeed, it would be

something of an oxymoron. These exams do not assess whether candidates, whatever their background, are able to manage this level of flexibility in consultations in settings of linguistic diversity. Standardised role-players tend to use a local British way of speaking and so monolingual UK graduates are not judged on whether they can carry out an effective consultation in linguistically challenging situations, with patients from different linguistic/cultural backgrounds. Such standardisation also means that IMGs, many of whom consult regularly in another expert language, have no opportunity to display this linguistic and cultural knowledge, which may be so valuable in everyday practice. In sum, simulations also raise questions about the fairness of assessments and how proportionate they are in an increasingly diverse society where there is no one right way of showing interpersonal effectiveness (Atkins *et al.* 2016).

5 Conclusion

The simulated consultation is a highly standardised genre, and in attempting to assess candidates' clinical skills it comes up against the untameable quality 'empathy'. Even if empathy were not assessed, the OSCE-type exam raises as many questions as it answers once we look outside the psychometrics of reliability and most types of validity. Since the institutional frame of these exams overrides any local relationship work, the consultation cannot be experienced as real and instead may depend, in part, on acting skills. Even though candidates can be trained to simulate the examined consultation, there are still several aspects of it which escape from the real. These include the powerful positioning of the role-player and the interactional evidence that they have opportunities to claim this power and asymmetry at a micro level (Atkins forthcoming). In addition, there is a lack of everyday contextual features such as the 'voice' of the computer in shaping the interaction and in eroding some of the modelled features of interpersonal skills, such as amount of eye contact (Swinglehurst *et al.* 2012).

Together with the tricky business of assessing the real through the unreal, the increasing focus on the soft interpersonal skills, and more especially empathy, raises serious questions about what empathy is, if it can be experienced and if so how, in a standardised exam setting. Here, relationship building over time and the deep values inherent in building capability (Fraser and Greenhalgh 2001; Bleakley 2003) are outsourced to an externally timed case where surface skills are voiced so that they can be monitored and assessed. And, to go one step further, there is the matter of whether 'true empathy' (Halpern 2003) is a necessary moral requirement of being a doctor at all.

Tied into these concerns is the issue of fairness. OSCE-style exams are talk-heavy, requiring more voiced empathy phrases and more interactional work to

inoculate them against sounding formulaic and insincere. Under the intense gaze of the examiners, a candidate whose style of communicating is likely to be somewhat different from theirs is under particular pressure. Small differences can have large consequences. And IMGs feel the weight of the water, which others may swim through more freely and lightly. At the same time, the diversity of communicative practices, typical of so many family and hospital practices today, are difficult to assess in such exams, again leading to concerns about whether things being equal means that they are fair.

Linguistic ethnography (LE), the approach used in the research illustrated here, drawing on sociolinguistics and anthropology (Copland *et al.* 2015), can also feel the weight of difference when used in medical settings. These methodologies can be viewed with scepticism by non-sociolinguists: since people talk and interact all the time, knowledge about language, communication and even culture can seem self-evident or simply uninteresting; also, there seems little space for such research in positivistic and psychometric paradigms where the notion of 'small is beautiful' (Schumacher 1973) is squeezed to the periphery. For sociolinguists, like IMGs, there is much hard discursive work to be done. We need to engage with the medical paradigm, which expects generalisations from big data, by tracing the source of small events and difference to wider social forces and theories through close contextual analysis (Burawoy 1998; Mitchell 1983; Small 2009). For example, we can analyse an intonation contour in the ever widening context of coded typical exam phrases, examiner feedback, interpersonal skills as designed into the exam, institutional systems of assessment, current medical discourses, performance and sociolinguistic theories of language use and difference, and feminised and equality discourses, as the examples above suggest.

A study such as this has value only if its practical relevance is clear and if it contributes, to whatever extent, to re-setting the terms of the debate about 'language' in OSCE-type exams. In the UK, this has involved us in long-term relationships with various institutions that teach and assess postgraduate medical professionals. There is, also, a wider debate around standardisation in increasingly pluralistic societies and how institutions address the challenges of face-to-face assessments in a globalised context. Chris Candlin in the 1970s was one of the first applied linguists to study the medical consultation when, in the UK, society was becoming increasingly diverse. And, more recently, with his associates, he was a significant voice in using close interactional analysis to enhance assessment and teaching practices (O'Grady and Candlin 2013). In the twenty-first century, this early work looks all too prescient, and the recent work increasingly timely, as we grapple with what it means to be professionally competent with a patient population from everywhere.

Appendix: Transcription conventions

CAN	Candidate's speech
RPL	Role-player's speech
PPP:	Pause
(0.8)	Pause timed to tenth of a second
(.)	Pause of less than (0.2) seconds
·hhhh	Inhalation
er::m	Extended word/sound
bi-	Unfinished word/sound
→	Level intonation
xxx	Inaudible sound
≈	Latched speech (first part)
+≈	Latched speech (second part)
[]	Overlapping speech
[]	

Acknowledgements

This work was supported by the Economic and Social Research Council [grant number ES/K00865X/1] to Sarah Atkins, University of Nottingham, 2013–16; and was originally supported by a Knowledge Transfer Partnerships award [grant number KTP008346] to Celia Roberts, King's College London, 2011–13. Ethical approval for the study was gained from King's College London, application REP(EM)/10/11-36.

About the authors

Sarah Atkins is Assistant Professor in Applied Linguistics at the University of Nottingham. Her research investigates language and professional communication in a range of workplace settings, with a particular focus on healthcare. She works on an ESRC-funded project at the University of Nottingham (ES/K00865X/1) looking at the use of medical simulations in communication skills training for postgraduate doctors. As a member of the university's *Linguistic Profiling for Professionals* unit, she also delivers workshops and training for practitioners, based on the findings of her applied linguistic research. Address for correspondence: Centre for Research in Applied Linguistics, School of English, Trent Building, University of Nottingham, NG7 2RD, UK. Email: Sarah.Atkins@nottingham.ac.uk

Celia Roberts is Professor Emerita in Sociolinguistics at King's College London. Her interests are in language and cultural processes in workplace and health contexts, their practical relevance to real-world situations of disadvantage and in research methods in ethnography and linguistics. Her publications include: *Language and Discrimination* (with Davies and Jupp, Routledge, 1992); *Achieving*

Understanding 1996 (with Bremer *et al.*); *Talk, Work and Institutional Order* (with Sarangi, De Gruyter, 1999); and *Performance Features in Clinical Skills Assessment* (with Atkins and Hawthorne, 2014). Address for correspondence: Centre for Language, Discourse and Communication, Franklin-Wilkins Building, Stamford Street, London, SE1 9NH, UK. Email: Celiaroberts11@gmail.com

References

- Atkins, S. (forthcoming) Assessing health professionals' communication through role-play: An interactional analysis of simulated versus actual GP consultations. *Discourse Studies*.
- Atkins, S., Roberts, C. and Hawthorne, K. (2016) Simulated consultations: A sociolinguistic perspective. *BioMedCentral Medical Education* 16 (16): 1–9. <https://doi.org/10.1186/s12909-016-0535-2>
- Balint, M. (1970) Quality, core values and the general practice consultation: Issues of definition, measurement and delivery. *Family Practice* 21 (4): 458–468.
- Bleakley, A. (2003) 'Good' and 'poor' communication in an OSCE: Education or training? *Medical Education* 37 (3): 186–187. <https://doi.org/10.1046/j.1365-2923.2003.01423.x>
- Blommaert, J. and Rampton, B. (2016) Language and superdiversity. In K. Arnaut, J. Blommaert, B. Rampton and M. Spotti (eds) *Language and Superdiversity*, 21–48. London: Routledge.
- Bonvicini, K., Perlin, M., Bylund, C., Carroll, G., Rouse, R. and Goldstein, M. (2009) Impact of communication training on physician expression of empathy in patient encounters. *Patient Education and Counseling* 75 (1): 3–10. <https://doi.org/10.1016/j.pec.2008.09.007>
- Boulet, J. R., Smee, S. M., Dillon, G. F. and Gimpel J. R. (2009) The use of standardized patient assessments for certification and licensure decisions. *Simulation in Healthcare* 4 (1): 35–42. <https://doi.org/10.1097/SIH.0b013e318182fc6c>
- Bouma, H. (2008) Is empathy necessary for the practice of 'good' medicine? *Open Ethics Journal* 2: 1–12. <https://doi.org/10.2174/1874761200802010001>
- Bourdieu, P. (1977) *Outline of a Theory of Practice*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511812507>
- Bradley, P. (2006) The history of simulation in medical education and possible future directions. *Medical Education* 40 (3): 254–262. <https://doi.org/10.1111/j.1365-2929.2006.02394.x>
- Brannick, M. T., Erol-Korkmaz, H. T. and Prewett, M. (2011) A systematic review of the reliability of objective structured clinical examination scores. *Medical Education* 45 (12): 1181–1189. <https://doi.org/10.1111/j.1365-2923.2011.04075.x>
- Brown, B., Crawford, P. and Carter, R. (2006) *Evidence-Based Health Communication*. Milton Keynes: Open University Press.
- Burawoy, M. (1998) The extended case method. *Sociological Theory* 16 (1): 4–33. <https://doi.org/10.1111/0735-2751.00040>

- Campion, P., Foulkes, J., Neighbour, R. and Tate, P. (2002) Patient-centredness in the MRCGP video examination: Analysis of a large cohort. *British Medical Journal* 325 (7366): 691–692. <https://doi.org/10.1136/bmj.325.7366.691>
- Cleland, J. A., Abe, K. and Rethans, J-J. (2009) The use of simulated patients in medical education: Association for Medical Education in Europe (AMEE) Guide No 42. *Medical Teacher* 31 (6): 477–486. <https://doi.org/10.1080/01421590903002821>
- Copland, F., Shaw, S. and Snell, J. (2015) *Linguistic Ethnography: Interdisciplinary Explorations*. Basingstoke, UK: Palgrave Macmillan. <https://doi.org/10.4135/9781473910607.n2>
- De la Croix, A. and Skelton, J. (2009) The reality of role-play: Interruptions and amount of talk in simulated consultations. *Medical Education* 43 (7): 695–703. <https://doi.org/10.1111/j.1365-2923.2009.03392.x>
- De la Croix, A. and Skelton, J. (2013) The simulation game: An analysis of interactions between students and simulated patients. *Medical Education* 47 (1): 49–58. <https://doi.org/10.1111/medu.12064>
- ELAN (2017) [Computer software, Version 5.0.0-beta]. Nijmegen: Max Planck Institute for Psycholinguistics. Available online: <https://tla.mpi.nl/tools/tla-tools/elan/>
- Erickson, F. and Schultz, G. (1982) *The Counselor as Gatekeeper*. New York: Academic Press.
- First, L. R., Chaudhry, H. J. and Melnick, D. E. (2013) Quality, cost, and value of clinical skills assessment. *New England Journal of Medicine* 368 (10): 963–964. <https://doi.org/10.1056/NEJMc1301008>
- Fraser, S. and Greenhalgh, T. (2001) Coping with complexity: Educating for capability. *British Medical Journal* 323 (7316): 799–803. <https://doi.org/10.1136/bmj.323.7316.799>
- Goffman, E. (1959) *The Presentation of Self in Everyday Life*. New York: Doubleday.
- Goffman, E. (1974) *Frame Analysis: An Essay on the Organization of Experience*. Cambridge, MA: Harvard University Press.
- Goffman, E. (2005) *Interaction Ritual: Essays in Face-to-Face Behavior*. London: Aldine Transaction Publishers.
- Greenhalgh, T. (2014) Future-proofing relationship-based care: A priority for general practice. *British Journal of General Practice* 64 (628): 580. <https://doi.org/10.3399/bjgp14X682357>
- Grugulis, I. and Vincent, S. (2009) Whose skill is it anyway? ‘Soft skills’ and polarization. *Work, Employment and Society* 23 (4): 597–615. <https://doi.org/10.1177/0950017009344862>
- Gumperz, J. (1982) *Discourse Strategies*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511611834>
- Halpern, J. (2003) What is clinical empathy? *Journal of General Internal Medicine* 18 (8): 670–674. <https://doi.org/10.1046/j.1525-1497.2003.21017.x>
- Hoekje, B. (2011) International medical graduate in U.S. higher education: An overview of issues for ESP and applied linguistics. In B. Hoekje and S. Tipton (eds) *English*

- Language and the Medical Profession: Instructing and Assessing the Communication Skills of International Physicians*, 3–20. Bingley, UK: Emerald Publishing.
- Hojat, M., Mangione, S., Nasca, T. J., Rattner, S., Erdmann, J. B., Gonnella, J. S. and Magee, M. (2004) An empirical study of decline in empathy in medical school. *Medical Education* 38 (9): 934–941. <https://doi.org/10.1111/j.1365-2929.2004.01911.x>
- Howie J., Heaney, D. and Maxwell, M. (2004) Quality, core values and the general practice consultation: Issues of definition, measurement and delivery. *Family Practice* 21 (4): 458–468. <https://doi.org/10.1093/fampra/cmh419>
- Iedema, R., Jorm, C., Braithwaite, J., Travaglia, J. and Lum, M. (2006) A root cause analysis of clinical error: Confronting the disjunction between formal rules and situated clinical activity. *Social Science and Medicine* 63 (5): 1201–1212. <https://doi.org/10.1016/j.socscimed.2006.03.035>
- Jamison, L. (2014) *The Empathy Exams: Essays*. Minneapolis, MI: Graywolf Press.
- Khan, K. Z., Ramachandran, S., Gaunt, K. and Pushkar, P. (2013) The objective structured clinical examination (OSCE): Association of Medical Education in Europe (AMEE) Guide No. 81. Part I: An historical and theoretical perspective. *Medical Teacher* 35 (9): e1437–e1446. <https://doi.org/10.3109/0142159X.2013.818634>
- Konijin, E. (2000) *Acting Emotions: Shaping Emotions on Stage*. Amsterdam: Amsterdam University Press. <https://doi.org/10.5117/9789053564448>
- Korkiakangas, T., Weldon, S.-M., Bezemer, J. and Kneebone, R. (2015) Video-supported simulation for interactions in the operating theatre (ViSIOT). *Clinical Simulation in Nursing* 11 (4): 203–207. <https://doi.org/10.1016/j.ecns.2015.01.006>
- Kurtz, S., Silverman, J. and Draper, J. (1998) *Teaching and Learning Communication Skills in Medicine*. Abingdon, UK: Radcliffe Medical Press.
- Larson, E. B. and Yao, X. (2005) Clinical empathy as emotional labor in the patient–physician relationship. *Journal of the American Medical Association* 293 (9): 1100–1106. <https://doi.org/10.1001/jama.293.9.1100>
- Lievens, F. and Sackett, P. R. (2012) The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *Journal of Applied Psychology* 97 (2): 460–468. <https://doi.org/10.1037/a0025741>
- MacWhinney, B. (2000) *The CHILDES Project: Tools for Analyzing Talk* (3rd edition). Mahwah, NJ: Lawrence Erlbaum Associates.
- Malhotra, A., Gregory, I., Darvill, E., Goble E., Pryce-Roberts, A., Lundberg, K., Konradsen, S. and Hafstad, H. (2009) Mind the gap: Learners’ perspectives on what they learn in communication compared to how they and others behave in the real world. *Patient Education and Counseling* 76 (3): 385–390. <https://doi.org/10.1016/j.pec.2009.07.024>
- Marshall, R. and Bleakley, A. (2009) The death of Hector: Pity in Homer, empathy in medical education. *Medical Humanities* 35 (1): 7–12. <https://doi.org/10.1136/jmh.2008.001081>
- McManus, I. C., Elder, A. T and Dacre, J. (2013) Investigating possible ethnicity and sex bias in clinical examiners: An analysis of data from the Membership of the Royal College

- of Physicians (MRCP(UK)) PACES and nPACES examinations. *BioMedCentral Medical Education* 13: 103. <https://doi.org/10.1186/1472-6920-13-103>
- Mitchell, J. (1983) Case and situational analysis. *Sociological Review* 31 (2): 187–211. <https://doi.org/10.1111/j.1467-954X.1983.tb00387.x>
- Neighbour, R. (1987) *The Inner Consultation – How to Develop an Effective and Intuitive Consulting Style*. Lancaster, UK: Kluwer Academic Press.
- Niements, N. (2013) From role-playing to role-taking: Interpreter's role(s) in healthcare. In C. Schäffner, Y. Fowler and K. Kredens (eds) *Interpreting in a Changing Landscape: Selected Papers from Critical Link*, 305–319. Amsterdam: John Benjamins.
- O'Grady, C. and Candlin, C. N. (2013) Engendering trust in a multiparty consultation involving an adolescent patient In C. N. Candlin and J. Crichton (eds) *Discourses of Trust*, 52–69. Basingstoke, UK: Palgrave Macmillan.
- Pedersen, R. (2009) Empirical research on empathy in medicine: Critical review. *Patient Education and Counseling* 76 (3): 307–322. <https://doi.org/10.1016/j.pec.2009.06.012>
- Roberts, C., Atkins, S. and Hawthorne, K. (2014) *Performance Features in Clinical Skills Assessment: Linguistic and Cultural Factors in the Membership Exam in the Royal College of General Practitioners*. London: King's College London with the University of Nottingham.
- Roberts, C. and Sarangi, S. (2003) Uptake of discourse research in interprofessional settings: Reporting from medical consultancy. *Applied Linguistics* 24 (3): 338–359. <https://doi.org/10.1093/applin/24.3.338>
- Roberts, C., Wass, V., Jones, R., Sarangi, S. and Gillett, A. (2003) A discourse analysis study of 'good' and 'poor' communication in an OSCE: A proposed new framework for teaching students. *Medical Education* 37 (3): 192–201. <https://doi.org/10.1046/j.1365-2923.2003.01443.x>
- Rubin, D. L., Healy, P., Gardiner, T., Zath, R. C. and Moore, C. P. (1997) Nonnative physicians as message sources: Effects of accent and ethnicity on patients: Responses to AIDS prevention counseling. *Health Communication* 9 (4): 351–368. https://doi.org/10.1207/s15327027hc0904_4
- Schumacher, E. F. (1973) *Small is Beautiful*. London: Blond and Briggs.
- Seale, C., Butler, C., Hutchby, I., Kinnersley, P. and Rollnick, S. (2007) Negotiating frame ambiguity: A study of simulated encounters in medical education. *Communication and Medicine* 4 (2): 177–187. <https://doi.org/10.1515/CAM.2007.021>
- Skelton, J. R. (2005) Everything you were afraid to ask about communication skills. *British Journal of General Practice* 55 (510): 40–46.
- Smajdor, A., Stöckl, A. and Salter, C. (2011) The limits of empathy: Problems in medical education and practice. *British Medical Journal of Medical Ethics* 37 (6): 325. <https://doi.org/10.1136/jme.2010.039628>
- Small, M. L. (2009) 'How many cases do I need?' On science and the logic of case selection in field-based research. *Ethnography* 10 (1): 5–38. <https://doi.org/10.1177/1466138108099586>

- Stepien, K. and Baernstein, A. (2006) Educating for empathy: A review. *Journal of General Internal Medicine* 21 (5): 524–530. <https://doi.org/10.1111/j.1525-1497.2006.00443.x>
- Stewart, M. (2001) Towards a global definition of patient-centred care. *British Medical Journal* 322 (7284): 444–445. <https://doi.org/10.1136/bmj.322.7284.444>
- Suchman, A. L., Markakis, K., Beckman, H. B. and Frankel, R. (1997) A model of empathic communication in the medical interview. *Journal of the American Medical Association* 277 (8): 678–682. <https://doi.org/10.1001/jama.1997.03540320082047>
- Swanson, D. B. and van der Vleuten, C. P. (2013) Assessment of clinical skills with standardized patients: State of the art revisited. *Teaching and Learning in Medicine: An International Journal* 25 (suppl. 1): S17–S25. <https://doi.org/10.1080/10401334.2013.842916>
- Swinglehurst, D., Greenhalgh, T. and Roberts, C. (2012) Computer templates in chronic disease management: Ethnographic case study in general practice. *British Medical Journal Open* 2 (6): e001754. <https://doi.org/10.1136/bmjopen-2012-001754>
- Thomassen, G. (2009) The role of role-play in medical education: Managing activity ambiguities in simulated consultations between medical students and patients. *Communication and Medicine* 6 (1): 83–93. <https://doi.org/10.1558/cam.v6i1.83>
- Vertovec, S. (2007) *New Dimensions of Complexity in Britain: Superdiversity, Transnationalism and Civil-Integration*. Think piece for the Commission on Integration and Cohesion, Department for Communities and Local Government. University of Oxford, Oxford. Available online: https://www.compas.ox.ac.uk/media/ER-2007-Complexities_Cohesion_Britain_CIC.pdf